

COGNITIVE LEVELS OF EVOLUTION: from pre-rational to meta-rational

Francis Heylighen*

PESP, Free University of Brussels, Pleinlaan 2, B-1050 Brussels, Belgium; *E-mail:* fheyligh@vnet3.vub.ac.be

ABSTRACT: The principle of natural selection is taken as a starting point for an analysis of evolutionary levels. Knowledge and values are conceived as vicarious selectors of actions from a repertoire. The concept of metasystem transition is derived from the law of requisite variety and the principle of hierarchy. It is defined as the increase of variety at the object level, accompanied by the emergence of a situation-dependent control at a metalevel. It produces a new level of evolution, with a much higher capacity for adaptation. The most important levels are discussed, with an emphasis on the level characterizing man as distinct from the animals. An analysis of the shortcomings of this "rational" system of cognition leads to a first sketch of how the next higher "meta-rational" level would look like.

1. INTRODUCTION

The *Principia Cybernetica Project* aims to develop an integrated philosophical system based on cybernetic and evolutionary concepts (Turchin, 1990). This system will be constructed collaboratively, with the support of computer technologies such as hypermedia, electronic mail and electronic publishing (Joslyn, 1990). A complete philosophy should contain at least an ontology or metaphysics, an epistemology, and an ethics or theory of values. Moreover, it should provide an answer to the fundamental questions: Who am I? Where do I come from? Where am I going to?

It is my contention that all these questions can be answered on the basis of an analysis of the general process of evolution, with a particular emphasis on the emergence of new levels of complexity or control. Indeed, the concept of evolution itself provides a foundation for a "process" metaphysics, like in the philosophies of Bergson, Teilhard and Whitehead, and for an "evolutionary" epistemology (Campbell, 1974). It allows a detailed analysis of the development of human intelligence from lower levels of evolution, thus answering the questions about who we are and where we come from. Finally, an extrapolation of the direction of evolution into the future may provide a first answer to the question of where we are going to, and thus provide a framework for a theory of values or ethics, aimed at the avoidance of "evolutionary dead ends".

In the present paper I will concentrate on the process which is at the base of the emergence of the higher, "cognitive" levels of evolution, starting with the origin of life. The lower—physical and chemical—levels will have to be addressed in another study.

2. DEFINING EVOLUTIONARY LEVELS

In this study, evolution will be assumed to be based on the principle of *blind variation and selective retention*. No teleology or guiding force will be assumed, except the

tautological principle that *stable* (intrinsically and with respect to the environment) systems tend to maintain, and hence will be "naturally" selected (Heylighen, 1989b), whereas unstable systems will be eliminated.

Each system is undergoing changes, either generated by internal processes, or by changes in the environment. Evolutionary selective systems will be those systems that "survive" the changes, i.e. that maintain a continuous identity, even though their state may have changed. We will suppose that the system is internally stable, i.e. that it is not destroyed by its own internal processes. In that case, the system will have to cope with external changes, which might perturb its internal organization, by adequate reactions. Such reactions are said to *compensate the perturbations* induced by the environmental changes, bringing the system back to its stable mode of operating (Heylighen, 1990c; Maturana & Varela, 1980).

Variety and selection

In general, different perturbations will require different reactions or compensations. This means that the larger the variety of potential perturbations, the larger the variety of compensations the system must be capable to execute. This can be understood from Ashby's (1958) *Law of Requisite Variety*. We will hence assume that every adaptive system disposes of a repertoire or variety of possible actions, that potentially compensate perturbations.

However, a repertoire alone is not very useful. The system must also be able to *choose* that action from the repertoire which is most likely to compensate a particular perturbation. Suppose the system does not have any specific criterion for making such decisions. In that case it will have to try out an action blindly or at random, and hope that it proves adequate. If the action is inadequate, the perturbation may destroy the system. Hence there will be a "natural" selection of adequately behaving systems.

For example, consider a species adapted to a specific environment. Suppose the environment changes, e.g. because the climate becomes colder. The species may adapt because mutations and recombinations of genes provide a variety of possible genomes, some of which are better adapted to a colder climate. Those genotypes will be selected by the environment and gradually replace the other genomes within the gene pool of the species. This process takes a lot of time, and many lives of individuals in the species are wasted because their (mutated) genes were inadequate. The process of adaptation would be much more efficient if the species would *know* which genomes would be adequate, and would be able to steer the mutations in the direction of such adapted genomes.

Cognition as vicarious selection

Knowledge can be defined as the ability to choose adequate actions from the repertoire, where "adequate" means: securing the survival of the system within its environment. Knowledge *selects* actions from the variety of potential actions, in the same way that natural selection selects by destroying inadequately behaving systems. The difference is that knowledge does not destroy actual systems, it only eliminates "potentialities". Knowledge substitutes for the environment, making selections before the environment is able to destroy the system. Knowledge can thus be defined as a substitute or "*vicarious*" selector, which internally represents—and thus allows the anticipation of—the selective action of the environment (see Campbell, 1974).

"Representation" here does not mean that the knowledge of an adaptive system consists of some kind of a homomorphic image or model of the environment. What is represented is not the *structure* of the environment, but its selective *function* or action, i.e. the interactions between system and environment which are potentially lethal. Knowledge in this sense is always to a large degree subjective, since it depends on the specific system which tries to survive within the environment.

Knowledge acts as a control on behavior. What distinguishes knowledge from control in the more traditional sense, is that cognitive control does not depend on a specific, fixed goal that the system is supposed to achieve. The only, very broadly defined goal is survival of the identity, but the identity itself may change during the lifetime of the system. Moreover, cognitive control is supposed to cope with a complex environment, and will in that way have to represent as much as possible features of that environment that are relevant to survival.

Hierarchy

Since the potential variety of the environment is infinite, the Law of Requisite Variety could be interpreted as saying that a large variety of actions is always more adaptive than a small one. Yet the problem is that a larger variety is more difficult to control: the larger the number of alternatives for action, the more difficult it is to make an informed decision. Hence organisms with a small variety may be more efficient in making an acceptable decision ("satisficing").

The only general method to increase the variety, while keeping the control manageable, seems to be *hierarchy*: factorizing the decision problem into different levels, such that a decision at the higher level constrains the variety of the decision at the lower level (Simon, 1962; Heylighen, 1989a). The decision at the higher level is easier to make since it only considers an abstract version of the repertoire of possible actions, characterized by a much lower variety than the concrete repertoire it represents. The decision at the lower level is easier because only a small part of the original variety remains after the higher level decisions have been made. In this way, the variety at each level can be kept acceptably small, while the variety of the entire system, which is the product of the varieties at the different levels, can be made arbitrarily large.

Let us examine how such a hierarchically structured system can emerge by variation and selection. Consider a number of simple systems, that each have just one control level. Now it is possible that those elements, through random combinations, would form an assembly that is more or less stable. That assembly hence will survive, and determine a new, higher-order system, with a continuous identity (Simon,

1962). However, in order to survive in an environment that is perturbing, the new system will again need a repertoire of actions, and a cognitive control for selecting those actions that will maintain the stable organization. The repertoire of the higher-order system automatically arises by considering all possible combinations of actions in the repertoire of the lower-order subsystems. The variety of the new repertoire is roughly equal to the product of the varieties of the subsystems. Each of the actions of the subsystems is locally controlled. However, for the global system, a new control is needed in order to coordinate the decisions made by the controlling subsystems. This control may again emerge through variation and selection.

Metasystem transitions

Such a process where a number of control systems are integrated into a single whole with the formation of a global control system is called a "metasystem transition" by Turchin (1977). The global control or vicarious selector appears itself through blind variation and selection. An important feature of the metasystem transition is that once the control at the higher or *meta*-level has emerged, the number of subsystems it controls will in general increase. Turchin calls this phenomenon the "branching growth of the penultimate level".

It can be understood on the basis of our analysis of variety: given an adequate selector for choosing actions, adaptivity increases with variety. Hence the emergence of an adequate control will mean that the increase of variety at the lower level becomes evolutionarily advantageous. Hence natural variation and selection will automatically lead to an increase in variety, until a point is reached where further increases would make an adequate control at the existing level too difficult. At that point the system is ready to undergo a new metasystem transition, with the emergence of a new control system, controlling the controls of the level below. Hence, the emergence of control leads to an increase of variety, whereas the increase of variety, if it is large enough, stimulates the emergence of a new control, in a positive feedback cycle.

The present analysis is in fact more general than Turchin's. Turchin assumes that the only way to increase variety is to physically multiply the number of subsystems, for example by replication of a basic template (e.g. a cell). I can easily imagine different ways to increase variety, by changing the mechanism by which the subsystems function and interact. For example, a fish that during evolution learns to swim faster, increases the variety of the different speeds it is capable to attain. This can be achieved for example by making the form of the fish more aerodynamic: there is no need for multiplication of subsystems. This is an example of a continuous increase of the range of a certain variable (here speed) that can be attained by the system. Another, discontinuous way to increase variety is to multiply the number of variables or degrees of freedom, by making fixed properties variable. Two examples, both discussed by Turchin, would be the capability of movement and the capability to learn for nerve cells. None of these transitions seems to require a multiplication of the number of subsystems: I do not see why the variability of synaptic connections would require the multiplication of cells, or of organelles.

In general it seems sufficient to remove certain (inadequate) constraints in order to increase the variability of the system. However, it is clear that variety cannot be increased indefinitely in this way. In the end each system has a maximum proper variety, determined by the product of the

degrees of freedom of each of its subsystems. If we want to increase the variety beyond this limit we will have to increase the number of subsystems, in accordance with Turchin's proposed mechanism. For example, it is clear that the variety of possible actions a human being is capable of cannot be implemented in a system with the size of a molecule, or even a single cell. Many different cells need to be integrated in order to provide the necessary variety. In conclusion, the Turchin mechanism of integration may well be the most important one for increasing variety, but it is not the most general one, and there are examples in evolution where it does not apply.

Values

The present definition of knowledge and control can in fact also be interpreted as a definition of goals or values. The "vicarious selector" can be interpreted as an evaluation criterion for actions, which provides a general direction for the behavior of the system. The ultimate goal or value in this analysis is of course survival, i.e. maintenance of the identity. If the survival to be achieved is understood to be of an indefinite—as long as possible—duration, then survival is equivalent to "immortality", as conceived by Turchin (1990). Both knowledge and more concrete goals and values can be derived from this ultimate value through the mechanisms of vicariousness and hierarchy. At the lower levels of evolution knowledge and values can in fact not be separated. It is only at a higher, "rational" level that the traditional dichotomy between "facts" and "values" becomes meaningful.

The main difficulty in founding an ethics based on the concept of survival is the question of which system is to survive: the individual, the society, the species, or the world ecology ("Gaia") as a whole? Ideally we would like all these systems to survive, but it is not obvious how to derive a more concrete value system which reconciles these different objectives in a practical and consistent way.

If we restrict ourselves to the survival of a single system then the concepts introduced in this section provide us with a first method to derive a more concrete value system. The goal of compensation of perturbations leads to a hierarchy of values, ordered according to the "urgency" of the perturbation: from the elimination of direct life-threatening disturbances, to the development of knowledge and skills that may be useful for survival in some far away and uncertain future (Heylighen, 1990d). The maximal development of this potential for adaptation may be called *self-actualization* (Heylighen, 1990d). In the present model it corresponds to the achievement of a maximum of controlled variety within the given control level. As such it corresponds to the last stage before the occurrence of a new metasystem transition. The objective of reaching a new metalevel would then be the natural prolongation of the goal of reaching self-actualization.

3. PRE-RATIONAL LEVELS

Origin of life

Let us now discuss the most important cognitive levels of evolution. The first level starts with the origin of life. In fact knowledge can be linked with life itself, since - in Maturana's phrase (Maturana & Varela, 1980) - "to live is to cognize". The living cell is characterized by a self-producing (autopoietic) organization, where the DNA controls the production of proteins and enzymes, and the enzymes control the production of DNA (Maturana & Varela, 1980). The

DNA can be seen as a structural template that selects chains of amino acids that form proteins needed for the proper functioning and survival of the cell. Hence it can be viewed as a vicarious selector (Campbell, 1974), embodying knowledge about how to survive within the given environment. This selection is situation-dependent, since different genes in the DNA may be active or inactive, depending on the presence or absence of certain enzymes, which itself depends on the overall chemical situation (boundary conditions) of the cell within its environment.

Control of position

The next level brings me to another criticism of Turchin's analysis. Turchin (1977) lists the emergence of the capability to *move* as the first of a series of metasystem transitions. Yet, in the present interpretation, a metasystem transition is assumed to consist of two separate, but interacting, processes: an increase of variety at the object level, and the emergence at the metalevel of a control mechanism, which selects from the variety at the level below in function of survival within a perturbing environment. In the case of movability, there only seems to be an increase of variety at the object level. There is no cognitive control (in the sense of situation-dependent selection). It is only in the next stage, which Turchin calls "irritability", that a control mechanism appears: the movement will now be a function of particular features of the environment sensed by the system. Sensory organs here act as media translating features of the environment into an internal representation (vicarious selector, see Campbell, 1974) which allows informed decision-making. Hence Turchin's transitions:

Movement = control of position, and
Irritability = control of movement,

should in my view be replaced by the single metasystem transition:

Movement + Irritability = control of position.

Irritability requires simple reflexes (Turchin, 1977). The control of simple reflexes leads to the level of complex reflexes, which we will not discuss here. We will rather immediately continue with the next, that of *learning*.

Learning

In the metasystem transition where learning emerges, it is less obvious how to separate the processes of variety increase and control emergence. The variety is here due to the fact that the synaptic strengths, which determine the probability that a stimulus or excitation would travel from one neuron to another one, are variable, so that the same pattern of excitation may lead to different results. The variability of synaptic connections accounts for what Turchin calls the capability to associate, i.e. to create variable associations between representations. Part of the control is realized through the Hebb rule, stating that a synaptic connection that is used often will increase its strengths so that the probability that it will be used later on increases. Hence the overall control of the flow of excitation in the neural network will depend on the sequence of situations experienced by the system, together with the initial organization, which determines the interpretation of those perceptions. The internal, inherited organization is what distinguishes between positive, pleasurable sensations (e.g. eating), and negative, painful ones (e.g.

being burned). The pleasure/pain evaluation will function as a second control, in addition to the Hebb rule, which will reinforce certain associations, while weakening other ones.

The Hebb rule makes it possible for the system to learn by experience. Excitatory patterns which are similar to patterns experienced in the past will be processed in a similar manner, making it possible to anticipate events following known events in a regular manner. However, the system is not capable to anticipate what will happen in situations that are essentially different from situations experienced in the past.

The associations that are formed through learning are limited to phenomena experienced in spatial or temporal contiguity. If when I see a dog I hear barking, I shall associate "barking" with "dog": the connections between the pattern of neural activation (primitive "concept") corresponding to the recognition of a barking sound and the one corresponding to the recognition of a dog will be strengthened. However, I shall normally not associate "dog" with "music", as long as I do not see a musical dog. In the case of a rabbit the association between "barking" and "dog" may be of vital importance: when the rabbit hears barking it knows that it must fly or hide in order not to be caught by the dog. However, it would be rather counterproductive if the rabbit would associate music with dog, and run away each time it hears the birds singing.

The reason why a rabbit cannot associate "dog" with "music" is that the primitive concept standing for "dog" or for "music" or for any other phenomenon that can be recognized by the cognitive system is *context-dependent*: it receives its meaning through the whole of associations it has with other primitive concepts. Those associations are the result of a slow process of learning based on many experiences of contiguous occurrence of two concepts. The concept (e.g. dog) cannot be separated from this context of associated concepts (e.g. barking) in order to bring it into contact with new concepts (e.g. music).

4. THE RATIONAL LEVEL

Separating concepts from their context

In the example sketched above, there are situations imaginable where the association between dog and music would nevertheless be useful. For example, a circus performer may gain a lot of money by training a dog for an act in which the dog would play a melody on a musical instrument. The fact that the human can imagine a dog producing musical sounds, while the rabbit cannot, signifies that the human has a larger variety of possible actions. Indeed once a situation is imagined, the human can find some way to realize it, e.g. by finding a smart dog and putting a lot of effort in training it. Even if we assume that a rabbit would be physically capable to train a dog, it is clear that the rabbit would never conceive this as a possible action, and hence would never look for adequate ways to execute it.

In a human cognitive system, on the other hand, concepts are separable from their context: they retain (part of) their meaning when brought into contact with radically different contexts. This can be understood by noticing that the concept can be distinguished from the concepts it is associated with, and that this distinction is stable or invariant: it does not change with the context (Heylighen, 1990c). Of course this context-independence is only partial or approximate: concepts still do have a lot of connotative (context-dependent) meanings, since they are rooted in an associative network

that learns through experience. This means that we shall have to postulate a specific mechanism that explains how any conceptual separation, however partial, is possible.

Symbols as support for concepts

Since the only physiological mechanism at our disposal is association through contiguity, a possible way to make concepts more independent is to associate each concept with some outside phenomenon that is physically independent from the other phenomena of its class, so that it can be easily combined with such phenomena without losing its identity or distinction. Such independent external phenomena that can be associated with a concept may be called *symbols*. A symbol is a stable, easily recognizable phenomenon that can be combined with other symbols. Each symbol denotes a concept. A combination of symbols hence represents a potential association of concepts.

The appearance of symbols is largely parallel with the appearance of the human *language*. Language consists of words (symbols, denoting concepts) that can be combined through syntactical rules in order to form sentences (denoting potential associations of concepts). A symbol represents a concept in a more abstract way, without most of its associations, so that there is room for associations with any of the other symbols. The system of symbols together with the rules for combining them may be called a *conceptual-symbolic code*.

The emergence of the rational level of cognition supported by a conceptual symbolic code is related to the emergence of language and culture as the system controlling social interactions. Indeed, communication through verbal language uses the same mechanism of symbols representing concepts in a stable, context-independent way. If the meaning of the symbol (word) would continuously vary with the context of associated concepts, no communication would be possible since a given symbol would have a different meaning for every individual who uses it, and no meaning could be transferred from one individual to another one by using such a symbol. However, there is no simple cause-and-effect relationship between language as a communication medium and rational cognition. Probably the relation is bicausal: the emergence of language would stimulate the emergence of rationality, since the exchange of meanings between different individuals would make it easier to abstract out those parts of the meaning which are invariant among the collectivity; on the other hand, the emergence of rationality would contribute to the emergence of language, since the internal stability of concepts would facilitate the exchange with other individuals (Heylighen, 1984).

Variety and control

Let us analyse this mechanism from the point of view of variety increase and control. The separability of concepts from the context leads to a large increase in the variety of possible associations, since now associations can be formed which were not experienced as spatial or temporal contiguities. This is yet another example of a mechanism for increasing variety, in addition to the multiplication of subsystems or variables, or the increase in the range of a variable. It consists in the partitioning of the system of associations, which in the previous level varies as a whole, into relatively independent subsystems (concepts) which can vary relative to each other, thus increasing the number of degrees of freedom of the system. Though the number of subsystems increases, we cannot really speak about a multiplication or replication

of subsystems, but rather about a splitting up of the original system.

This increase in variety necessitates a new control mechanism, guiding the selection of concept associations. This control consists of different features. There is first a set of *generative rules*, which in the case of language is called grammar or syntax, and which constrains the combinations of concepts by distinguishing different categories of concepts (nouns, verbs, adjectives, ...), which must be combined in specific ways. Second, there still are the control mechanisms of the lower level of experiential cognition: the evaluation of combinations of concepts by the associations they have with other concepts, and their positive or negative appraisals. Third, there are a number of conceptual-symbolic representations of selection criteria: goals and values. These are representations of features of possible future states that are to be desired or to be avoided. Fourth, there are a number of rules of thought based on culture: tradition, religion, philosophy, logic, science...

Rationality, free will, self-consciousness

This mechanism allowing the controlled recombination of concepts makes it possible to explain all the typical characteristics of human intelligence, as distinguished from animal intelligence. The fact that different combinations of concepts can be conceived without experiencing them in contiguity, can be interpreted as a basic form of imagination or *creativity*. The fact that there is a control mechanism allowing the choice between a set of alternatives conceived in this way, can be understood as the basis for *free will*. The fact that a controlled sequence of combinations can be generated and explored as to its consequences may be called *rationality* or the capacity for thinking. The fact that an individual can have a concept of himself, and can combine it with other concepts in order to think about himself and compare himself with others can be called *self-consciousness*. By combining the concepts of situations he wants to achieve with the concepts of things he has at his disposal he may conceive ways to use these things as *tools*.

5. TOWARDS A META-RATIONAL LEVEL

Shortcomings of rational cognition

The rational level of cognition is evidently not perfect, and we may expect that still higher levels will emerge during the course of evolution. In order to get an idea of the direction in which that evolution would proceed, it is interesting to analyse the shortcomings of rationality. The rational mechanism requires the splitting up of the continuous field of experience embodied in the network of neural associations into a discrete set of stable concepts. Evidently, an important part of the (connotative) meaning contained in the associations is lost in that way. We might keep more of the meaning by making the set of concepts larger. However, this will increase the variety in an unlimited way, since there is an infinity of ways to segment a continuous field into a discrete set of distinct concepts. The larger the set of concepts, the larger the set of possible combinations of concepts. The control mechanisms needed for choosing adequate combinations out of that diversity of potential combinations are obviously limited, and beyond a certain level of complexity they will become ineffective as guides for decision-making.

The set of concepts used by an individual is normally provided by the culture in which he lives. Since the number of possible concepts is infinite, the set of actual concepts will

always be to a certain degree arbitrary. When the individual comes into contact with a different culture he will encounter concepts which do not fit into his conceptual code. This will make it difficult to understand expressions from that culture. The only way to grasp the meanings from that other culture is by using the rather slow mechanism of associative learning: the conceptual code itself does not provide any means to extend its range of concepts. Moreover, since, as we remarked before, the separation of a concept from its context is never complete, there will remain certain connotative meanings to each concept which are difficult to communicate because of their subjectivity. Again the conceptual code does not provide any means to tackle this problem. The result is that individuals will never understand each other completely even though they do speak the same language, and it will be very difficult for them to correct the misunderstandings using that language. In short, each conceptual code will be impaired by its (unadmitted) subjectivity, by its limited number of concepts, and by its tendency to reduce holistic phenomena to combinations of discrete elements.

If we look at cultural evolution we can distinguish different movements which have tried to somehow transcend these limitations of the conceptual code (Heylighen, 1984). For example, *science* has attempted to eliminate the subjectivity of conventional languages by grounding the meaning of terms through formalization and operationalization (Heylighen, 1990a). *Art* attempts to overcome the limited expression of connotative meanings by finding alternative, non-reductionistic ways to symbolize associative, experiential meanings. *Religion* tries to transcend the fragmentation of the continuous field of experience by proposing an overarching system of faith that binds the different domains of knowledge and experience together, and by stimulating certain non-rational forms of consciousness (mystical experiences) in which the fragmentation of experience is abolished. *Philosophy* similarly tries to integrate different domains of experience but does this by analysing the meanings of terms and the limitations imposed by conventional language. As such it stands in between science and religion.

Acceleration of socio-cultural evolution

A more detailed look at cultural history makes us distinguish two stages in these attempts to transcend the rational level of cognition (Heylighen, 1984). The first, "classical" stage was not very successful, since the origin of the intuitively felt limitations was not really understood. It is only around the beginning of the 20th century, that a first awareness arose of the impossibility to transcend the limitations within the conceptual code itself. This awareness can be associated with cultural revolutions such as conventionalism and formalism in mathematics, relativity and quantum indeterminacy in physics, the discovery of the unconscious in psychology, surrealism, experimentation and abstraction in art, and analysis of the role of language in philosophy (Wittgenstein).

It led to a release of most of the constraints (taboos, dogmas, esthetic and epistemic principles, ...) that still restricted the creation of new conceptual or associative systems outside the existing conceptual codes. Together with the creation of better communication media, this resulted in an ever faster flux of new concepts, knowledge, ideas, fashions, codes, models, languages, ... in the most diverse domains of culture: science, technology, management, politics, religion, art, fashion, mass media, ... This ever accelerating wave of cognitive and social change has been well described by Toffler (1970). In his analysis he stressed the danger resulting from

the fact that at a certain level of change our cognitive systems is no longer capable to cope with the complexity. The whole of perturbations of our physical and psychical equilibrium resulting from this forms a new type of disease which Toffler called "future shock".

The appearance of such a syndrome can be easily understood from our observation that if the variety of possible actions is larger than the capacity for control of the cognitive system, the system is ill-adapted. The larger the number of alternatives to be considered, the higher the stress imposed by decision-making. Paradoxically, though our power to solve problems is greater than ever before, the problems also appear more complex than ever before, and we may feel even more helpless than ever before (Heylighen, 1989a). The only real solution to this problem is to introduce a new level of control, i.e. to undergo a metasystem transition. This transition will bring us to a higher, "meta-rational" level of cognition.

Meta-rational consciousness

In order to understand how the new control mechanism will function we must analyse just in which way variety is increased. The overall phenomenon seems to be an exponential increase of the available rational knowledge: concepts and rules for combining concepts. However since the rational knowledge cannot be separated from the associative, experiential knowledge in which it is rooted, this latter form of knowledge must increase too.

By increase of knowledge I do not mean just an increase in the number of facts or data, but more importantly an increase in the number of conceptual systems (codes, models, representations, languages) which are used to dissect reality and to represent it in a manageable form. The same reality can be dissected in an infinite numbers of ways, leading to incompatible or irreducible representations. So the first requirement for attaining a meta-rational level of consciousness is the awareness that conceptual systems are relative, that there is no one true representation of the world, but that there are an infinity of complementary representations which each have their proper advantages and disadvantages.

Such awareness is not sufficient, however. We also need practical rules for guiding us in the search for that conceptual system which is most appropriate for the problem we want to solve. The "problem" is defined objectively by the state of the environment (available resources and constraints) and subjectively by the goal we want to achieve. The analysis of theory- and model-building in different sciences has already led to a lot of knowledge that can be used for finding adequate representations. In addition to that we need an increased understanding of the basic dynamics of evolution, so that we can better understand our own role in the on-going transitions, and choose our own long-term plans and goals in a more conscious fashion.

Knowledge about model-building, and about goal-setting, based on a system of values derived from evolutionary theory, provides us with the first elements of the control we need to cope with the increased socio-cultural variety. What remains to be done is to integrate that "metacognitive" knowledge in the form of a universal framework: an *adaptive metarepresentation* (Heylighen, 1988, 1990a). I will here not go into further detail about my own research in this direction. It will suffice to say that my approach is based on a "dynamics of distinctions", which provides general principles for the creation of new higher-order distinctions (elementary concepts), by analysing conceptual networks

with the aid of the mathematical concept of relational closure (Heylighen, 1989a,b, 1990c).

Technological support for cognition

Individual awareness and (meta)knowledge are still not sufficient to control the huge amount of new data, concepts and rules. We also need external support for storing, accessing and processing that information. We cannot wait until our brain has by natural selection evolved into a system that is better adapted to process such complex data in a fast and reliable way. There is no reason to despair, though: the most spectacular evolution of the last decades is the emergence of increasingly powerful technologies for storing, transmitting and processing information. The increase in speed and storage of computers and related media seems exponential. Hence it is clear that we will not have to wait very long to see systems emerge whose capacities will surpass anything that has been imagined before.

The main challenge will be to integrate this technology with the metacognitive frameworks and with our existing information-processing capacities (sensory organs and brain). If the metarepresentational framework is formulated in a more or less formal way, it will not be difficult to implement it on existing computers. On the other hand, existing computer systems are not very well suited for a direct and intuitive connection with the human mind. What we need is a much more profound development of man-machine interfaces, taking into account the associative, experiential, context-dependent basis for the representation of knowledge in the brain (Heylighen, 1990c).

Existing computers are modelled after the rational-conceptual mode of cognition: they work with discrete chunks and fixed rules for their combination. Although in the human mind the rational mechanism is implemented on a higher control level than the associative mechanism, this does not mean that it is intrinsically superior. Its full power only emerges by its being rooted in an associative network, which complements it by providing the necessary variety. That is why the human brain is still much more powerful than existing computers functioning through a purely rational mechanism, without the power to form associations by experience.

Some recently emerging paradigms promise to bridge the gap between neural-associative and rational mechanisms of cognition (Heylighen, 1990c): 1) *connectionism* is a paradigm that allows computers to simulate the associative learning typical of the human brain at the most primitive level; 2) *hypermedia* provide a way to easily translate existing associative knowledge of a higher level in a form that can be easily stored by computers; 3) *cyberspace* or *virtual reality* is a philosophy for building man-machine interfaces which takes into account the full perceptual-cognitive mechanism a human uses to cope with the ordinary three dimensional world of movable objects and shapes. A further integration of those three approaches would lead us to imagine a direct connection between brain activities and processes inside a computer, for example by means of a sophisticated device for registering and inducing brain waves. In that case, the virtual reality of the representation would be generated directly, in the same way as a dream, i.e. without any intervening sensory stimuli.

Superbeings vs. Metabeings

Let us conclude by trying to imagine how the appearance of this new, meta-rational level of cognition would affect us as

human beings. In Turchin's (1977, 1990) view, the next meta-system transition would lead to the emergence of what he calls a human *superbeing* resulting from the integration of different individuals in a collective system. Such an integration would be similar to the way individual cells were integrated to form multicellular organisms. It would occur through much more direct communication channels connecting the nervous systems of the individuals. This view of the future follows logically from Turchin's definition of a meta-system transition as the integration of subsystems together with the emergence of a control level. Yet it leaves him with the fundamental contradiction between integration and freedom, since the social integration of individuals will suppress part of the individual's freedom, whereas Turchin acknowledges creative freedom to be the essence of life (Turchin, 1990; Turchin & Joslyn, 1990).

In the present interpretation, on the other hand, there is no such contradiction between freedom and control. Indeed, a metasystem transition is here defined as an increase of variety (and hence of freedom) at the object level, together with the emergence of a control at the metalevel which allows making informed, situation-dependent choices between the variety of actions available at the object level. Each metasystem transition hence by definition increases the overall freedom of the system. There is no difficulty with the constraints imposed upon the subsystem by integration since integration of multiple subsystems is not necessary to increase variety: other ways to achieve variety have been shown to exist.

In the present view, the newly emerging control would be situated on the level of the individual rather than on the level of society. Each human individual would dispose of a metarepresentational framework, implemented through an advanced man-machine interface, that would help him or her in manipulating knowledge, in creating new concepts and theories, and in efficiently gathering and organizing all the existing facts and values (s)he needs to solve his/her problems. Practically, what was considered to be a privilege that only a few persons of genius might achieve during their lifetime, namely the creation of a completely new theory modelling important parts of reality, would now become an automatic, everyday activity. This construction of a model would come as natural to meta-humans as the formation of a sentence in verbal language comes to us.

Of course, the existence of such a metacognitive framework and its accompanying communication and information technology would drastically enhance the possibilities for communication between different individuals. At least it would allow overcoming the pitfalls of (unadmitted) subjectivity and reductionism associated with the conceptual code defining the rational mechanism of cognition. Hence the metarational system would make a further integration of individuals certainly easier.

However, the question is whether such a further integration would be evolutionarily advantageous, given that each individual would have direct access to the whole of knowledge available to humanity, in order to design his own models, and solve his own problems. Clearly, like in each collectivity or society there must be certain rules that govern the whole and that restrain individuals in their actions that might endanger the survival of the other individuals or the group (e.g. war, pollution, ...). For example, in most species of animals there is a taboo against killing members of the same species. Yet this does not mean that there is an advantage in integrating the species as a whole into a superanimal. Too strong integration of the individuals would make the species

more vulnerable, since a highly integrated system (e.g. a multicellular organism) can be killed by killing a few of its subsystems (e.g. cells from the heart).

The main purpose of integration is, as Turchin remarks, to increase (controlled) variety, but it can also lead to a diminishing of the uncontrolled variety realized by the diversity between the individual subsystems. If the second type of variety becomes too low, a single cause might kill the whole of the subsystems. A well-known example is the danger of agriculture based on genetically identical or similar plants: a single disease or parasite may kill the whole culture. In conclusion, there seems to be a trade-off between variety gained through integration and variety lost through integration, which makes that the optimum level of survival will be achieved neither by complete integration nor by complete independence of subsystems. Where that optimum level of integration will be situated after the next metasystem transition is a question which at this stage seems difficult to answer.

ACKNOWLEDGMENTS:

My views on these issues have profited much from recent discussions over electronic mail with Valentin Turchin and Cliff Joslyn.

REFERENCES

1. Ashby W.R. (1958): "Requisite Variety and Implications for Control of Complex Systems", *Cybernetica* 1, p. 83-99.
2. Campbell D.T. (1974): "Evolutionary Epistemology", in: *The Philosophy of Karl Popper*, Schilpp P.A. (ed.), (Open Court Publish., La Salle, Ill.), p. 413.
3. Heylighen F. (1984): "Transcending the Conceptual-Symbolic Code", "*O*": *tijdschrift voor filosofische beproevingen* 7, p. 119 - 142. [in Dutch]
4. Heylighen F. (1988): "Formulating the Problem of Problem-Formulation", in: *Cybernetics and Systems '88*, Trappel R. (ed.), (Kluwer Academic Publishers, Dordrecht), p. 949-957.
5. Heylighen F. (1989a): "Coping with Complexity: concepts and principles for a support system", in: *Proceedings of the Int. Conference "Support, Society and Culture: Mutual Uses of Cybernetics and Science"*, Glanville R. & de Zeeuw G. (eds.), (IWA, University of Amsterdam), p. 26-41.
6. Heylighen F. (1989b): "Self-Organization, Emergence and the Architecture of Complexity", in: *Proc. 1st European Cong. on System Science* (AFCET, Paris), p. 23-32.
7. Heylighen F. (1990a): *Representation and Change. A Metarepresentational Framework for the Foundations of Physical and Cognitive Science*, (Communication & Cognition, Gent).
8. Heylighen F. (1990b): "Autonomy and Cognition as the Maintenance and Processing of Distinctions", in: *Self-Steering and Cognition in Complex Systems*, Heylighen F., Rosseel E. & Demeyere F. (eds.), (Gordon and Breach, New York), p. 89-106.
9. Heylighen F. (1990c): "Design of an Interactive Hypermedia Interface Translating between Associative and Formal Problem Representations", *International Journal of Man-Machine Studies* (in press).

10. Heylighen F. (1990d): "A Cognitive-Systemic Reconstruction of Maslow's Theory of Self-Actualization", submitted to *Systems Research*.
11. Joslyn C. (1990): "Tools for the Development of Consensually-Based Philosophical Systems: a feasibility study for the Principia Cybernetica Project", in: *Proc. 8th Int. Conf. of Cybernetics and Systems*, C.N. Manikopoulos (ed.) (this volume).
12. Maturana H. and Varela F. (1980): *Autopoiesis and Cognition: the realization of the living*, (Reidel, Dordrecht).
13. Simon H.A. (1962): "The Architecture of Complexity", *Proceedings of the American Philosophical Society* 106, p. 467-482.
14. Toffler A. (1970): *Future Shock*, (Random House, New York).
15. Turchin V. (1977): *The Phenomenon of Science*, (Columbia University Press, New York).
16. Turchin V. (1990): Cybernetics and Philosophy, in: *Proc. 8th Int. Conf. of Cybernetics and Systems*, C.N. Manikopoulos (ed.) (this volume).
17. Turchin V. and Joslyn C.: (1990) "The Cybernetic Manifesto", *Kybernetes* 19:2-3.

* Senior Research Assistant NFWO
(Belgian National Fund for Scientific Research)